On the weakest constraint qualification for strong local minimizers

Oliver Stein* Maximilian Volk#

June 23, 2022

Abstract

The strong local minimality of feasible points of nonlinear optimization problems is known to possess a characterization by a strengthened version of the Karush-Kuhn-Tucker conditions, as long as the Mangasarian-Fromovitz constraint qualification holds. This strengthened condition is not easy to check algorithmically since it involves the topological interior of some set. In this paper we derive an algorithmically tractable version of this condition, called strong Karush-Kuhn-Tucker condition, and we show that the weakest condition under which a feasible point is a strong Karush-Kuhn-Tucker point for every at this point continuously differentiable objective function possessing the point as a strong local minimizer, is the Guignard constraint qualification.

Keywords: Strong local minimizer, strict local minimizer of order one, Guignard constraint qualification, weakest constraint qualification

Preliminary citation: Optimization Online, Preprint ID 2022-06-8962, 2022

1 Introduction

We consider nonlinear optimization problems of the form

$$P: \quad \min_x f(x) \quad \text{s.t.} \quad g(x) \le 0,\ h(x) = 0$$

with defining functions $f \in C^1(\mathbb{R}^n, \mathbb{R})$, $g \in C^1(\mathbb{R}^n, \mathbb{R}^p)$ and $h \in C^1(\mathbb{R}^n, \mathbb{R}^q)$. The feasible set of P will be denoted by X, $\nabla f(x)$ stands for the gradient of f at x, and $\nabla g(x)$, $\nabla h(x)$ are the Jacobians of g and h, respectively, at x. With the active index set $A(\bar{x}) = \{i \in \{1, \dots, p\} \mid g_i(\bar{x}) = 0\}$ of $\bar{x}$ the matrix $\nabla g_A(\bar{x})$ possesses the columns $\nabla g_i(\bar{x})$, $i \in A(\bar{x})$.

*Institute of Operations Research (IOR), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, stein@kit.edu

#Institute of Operations Research (IOR), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany, maximilian.volk@student.kit.edu

We will be interested in necessary and sufficient optimality conditions for strong local minimizers of P, that is, points $\bar{x} \in X$ for which a neighborhood U and some $\alpha > 0$ exist with

$$\forall\, x \in X \cap U: \quad f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|.$$

Strong local minimizers are also called strict local minimizers of order one. Since strong local minimizers $\bar{x}$ of P are local minimizers, under some constraint qualification they are necessarily Karush-Kuhn-Tucker (KKT) points, that is, there exist $\lambda \in \mathbb{R}^p$ and $\mu \in \mathbb{R}^q$ with

$$\nabla f(\bar{x}) + \nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\mu = 0,\ \lambda \geq 0. \tag{1}$$

Moreover, since strong local minimizers are special local minimizers, one may expect that they also satisfy a strengthened version of the KKT conditions, and that this condition may even be sufficient for strong local minimality. Such a condition is given in [7, Th. 3.6] under the Mangasarian-Fromovitz constraint qualification (MFCQ) at $\bar{x}$, which assumes $\operatorname{rank} \nabla h(\bar{x}) = q$ and the existence of some vector $d \in \mathbb{R}^n$ with $\nabla g_A(\bar{x})^\intercal d < 0$ and $\nabla h(\bar{x})^\intercal d = 0$. Note that (1) may be rewritten as

$$-\nabla f(\bar{x}) \in \{\nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\mu \mid \lambda \in \mathbb{R}^p,\ \mu \in \mathbb{R}^q,\ \lambda \geq 0\}.$$

Theorem 1.1 ([7]). *Let the MFCQ hold at $\bar{x} \in X$. Then $\bar{x}$ is a strong local minimizer of P if and only if*

$$-\nabla f(\bar{x}) \in \operatorname{int}\{\nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\mu \mid \lambda \in \mathbb{R}^p,\ \mu \in \mathbb{R}^q,\ \lambda \geq 0\} \tag{2}$$

holds.

The aim of this paper is twofold. Firstly, we will derive an algorithmically more tractable version of the condition (2) and, secondly, we wish to identify a weakest constraint qualification under which strong local minimality is characterized by (2). Both is possible by using techniques which were introduced for the characterization of strict local Pareto optimal points of order one in multicriteria optimization [1, 3]. As, to the best of the authors' knowledge, the corresponding results in single objective optimization have not been formulated so far, the present paper first closes this gap and then studies the mentioned weakest constraint qualification.

2 A stationarity condition

For the following result we define the set $C_\leq(f, \bar{x}) = \{d \in \mathbb{R}^n \mid \nabla f(\bar{x})^\intercal d \leq 0\}$ of (potential) descent directions for f at $\bar{x} \in X$ and the tangent cone

$$T(X, \bar{x}) = \{d \in \mathbb{R}^n \mid \exists\, t^k \searrow 0,\ (x^k) \subseteq X : \lim_k (x^k - \bar{x})/t^k = d\}$$

to X at $\bar{x}$. The proof of the following result employs similar ideas as the ones of [3, Th. 4.1] and [7, Th. 3.2].

Lemma 2.1. *A point $\bar{x}$ is a strong local minimizer of P if and only if $\bar{x} \in X$ and $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x}) = \{0\}$ hold.*

Proof. Let $\bar{x} \in X$ not be a strong local minimizer of P. Then for each $k \in \mathbb{N}$ there exists some $x^k \in X$ with $\|x^k - \bar{x}\| \leq 1/k$ and $f(x^k) < f(\bar{x}) + (1/k)\|x^k - \bar{x}\|$. The sequence (t^k) with $t^k = \|x^k - \bar{x}\|$ satisfies $t^k \searrow 0$ and, by the compactness of the unit sphere, without loss of generality the sequence of directions $d^k = (x^k - \bar{x})/t^k$ converges to some $d \in T(X, \bar{x})$ with $\|d\| = 1$. Moreover, the differentiability of f yields $\nabla f(\bar{x})^\intercal d = \lim_k (f(x^k) - f(\bar{x}))/t^k$ with $(f(x^k) - f(\bar{x}))/t^k < 1/k$ for all k and, thus, $\nabla f(\bar{x})^\intercal d \leq 0$. This means that $d \neq 0$ lies in $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x})$ and therefore $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x}) \supsetneq \{0\}$ holds.

On the other hand, let $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x}) \supsetneq \{0\}$ for $\bar{x} \in X$ and choose some $d \neq 0$ from $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x})$. Then there exist some $t^k \searrow 0$ and $(x^k) \subseteq X$ with $d^k = (x^k - \bar{x})/t^k \to d$. Assume that $\bar{x}$ is a strong local minimizer. Then, with some $\alpha > 0$, for all sufficiently large k we have $(f(x^k) - f(\bar{x}))/t^k \geq (\alpha\|x^k - \bar{x}\|)/t^k = \alpha\|d^k\|$. This yields the contradiction $0 \geq \nabla f(\bar{x})^\intercal d = \lim_k (f(x^k) - f(\bar{x}))/t^k \geq \alpha\|d\| > 0$. □

Example 2.2. *For $n = 2$ let $f(x) = x_1 + x_2$, $g_1(x) = -x_1$, $g_2(x) = -x_2$ and $g_3(x) = x_1 x_2$. Then at $\bar{x} = 0 \in X$ we observe $C_{\leq}(f, \bar{x}) = \{d \in \mathbb{R}^2 \mid d_1 + d_2 \leq 0\}$ and $T(X, \bar{x}) = X = (\mathbb{R}_{\geq} \times \{0\}) \cup (\{0\} \times \mathbb{R}_{\geq})$. Therefore $C_{\leq}(f, \bar{x}) \cap T(X, \bar{x}) = \{0\}$ holds and $\bar{x}$ is a strong local minimizer.*

3 Strong Karush-Kuhn-Tucker points

In the sequel the following notion will be useful.

Definition 3.1. *We call $\bar{x} \in X$ a* strong Karush-Kuhn-Tucker point *if*

$$\operatorname{rank}(\nabla f(\bar{x}), \nabla g_A(\bar{x}), \nabla h(\bar{x})) = n \tag{3}$$

holds and if there exist $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^q$ with

$$\nabla f(\bar{x}) + \nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\mu = 0, \;\; \lambda > 0. \tag{4}$$

We remark that, under (4), the condition (3) is equivalent to the rank n of any matrix resulting from $(\nabla f(\bar{x}), \nabla g_A(\bar{x}), \nabla h(\bar{x}))$ by deletion of one of the columns $\nabla f(\bar{x})$, $\nabla g_i(\bar{x})$, $i \in A(\bar{x})$. In particular, the linear independence constraint qualification (LICQ) $\operatorname{rank}(\nabla g_A(\bar{x}), \nabla h(\bar{x})) = |A(\bar{x})| + q$, the identity $|A(\bar{x})| + q = n$ and the strict complementary slackness condition $\lambda > 0$ are sufficient for $\bar{x} \in X$ to be a strong KKT point. These conditions are, however, not necessary since at strong KKT points $|A(\bar{x})| + q > n$ may hold, and the multipliers λ and μ do not need to be unique.

Example 3.2. *In Example 2.2 we have $\nabla f(\bar{x}) = (1, 1)^\intercal$, $A(\bar{x}) = \{1, 2, 3\}$ and $\nabla g_A(\bar{x}) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \end{pmatrix}$ so that LICQ is violated at $\bar{x}$ while (3) is satisfied. Indeed, $\bar{x}$ is even a strong KKT point with the (nonunique) multiplier $\lambda = (1, 1, 1)^\intercal > 0$.*

We will characterize strong KKT points by means of Tucker's theorem of the alternative.

Lemma 3.3 ([4])**.** *For matrices A and B, with A being nonvacuous, exactly one of the following alternatives hold:*

a) $Ax \le 0$, $Ax \ne 0$, $Bx = 0$ possesses a solution x.

b) $A^\intercal y + B^\intercal z = 0$, $y > 0$ possesses a solution (y, z).

The proof of the following characterization is identical to the one of [1, Th. 3.4] where, however, a weaker assertion is stated as the result of the proof. For completeness we repeat the arguments here. The set $L(g, h, \bar{x}) = \{d \in \mathbb{R}^n \mid \nabla g_A(\bar{x})^\intercal d \le 0, \nabla h(\bar{x})^\intercal d = 0\}$ is the linearization cone to X at $\bar{x}$.

Lemma 3.4. *A point $\bar{x}$ is a strong KKT point of P if and only if $\bar{x} \in X$ and $C_{\le}(f, \bar{x}) \cap L(g, h, \bar{x}) = \{0\}$ hold.*

Proof. With $A^\intercal = (\nabla f(\bar{x}), \nabla g_A(\bar{x}))$ and $B^\intercal = \nabla h(\bar{x})$ we have $C_{\le}(f, \bar{x}) \cap L(g, h, \bar{x}) = \{0\}$ if and only if the system $Ad \le 0$, $Bd = 0$ possesses only the trivial solution $d = 0$. The latter is equivalent to the fact that, both, the system $Ad = 0$, $Bd = 0$, $d \ne 0$ is unsolvable, and the system $Ad \le 0$, $Ad \ne 0$, $Bd = 0$ is unsolvable. The unsolvability of the first system is equivalent to the linear independence of the n rows of the matrix $(A^\intercal, B^\intercal)$, that is, to (3). Moreover, by Lemma 3.3 the second system is unsolvable if and only if there exist $\kappa > 0$, $\lambda > 0$ with $\kappa \nabla f(\bar{x}) + \nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\mu = 0$. After division of this equation by κ (and renaming λ) this is condition (4). □

As the relation $T(X, \bar{x}) \subseteq L(g, h, \bar{x})$ is true without further assumptions, the combination of Lemma 2.1 and Lemma 3.4 yields that being a strong KKT point is sufficient for $\bar{x} \in X$ to be a strong local minimizer. If additionally the Abadie constraint qualification (ACQ) $L(g, h, \bar{x}) \subseteq T(X, \bar{x})$ holds at $\bar{x}$, then the same combination implies that being a strong KKT point is necessary for $\bar{x} \in X$ to be a strong local minimizer.

Since the ACQ is weaker than the MFCQ at $\bar{x}$ we particularly obtain that under the MFCQ a point $\bar{x}$ is a strong local minimizer if and only if it is a strong KKT point. This means that condition (2) in Theorem 1.1 is equivalent to $\bar{x} \in X$ being a strong KKT point, where the strong KKT property is easier to check algorithmically than the topological condition (2).

4 The Guignard constraint qualification

As we have seen in the previous section, strong local minimality of a point $\bar{x} \in X$ is characterized by the strong KKT conditions not only under MFCQ, but also under the weaker ACQ. This raises the question if one can identify a weakest constraint qualification for this setting. In fact, in Example 2.2 the ACQ is violated at $\bar{x}$, but $\bar{x}$ is still, both, a strong local minimizer and a strong KKT point.

Recall from [2] that the Guignard constraint qualification $T^\circ(X, \bar{x}) \subseteq L^\circ(g, h, \bar{x})$ is the weakest condition under which a point $\bar{x} \in X$ is a KKT point for every at $\bar{x}$ continuously differentiable function f possessing $\bar{x}$ as a local minimizer on X. Here $A^\circ = \{v \in \mathbb{R}^n \mid v^\intercal d \le 0 \, \forall d \in A\}$ denotes the polar cone of a cone $A \subseteq \mathbb{R}^n$. It is known that $T^\circ(X, \bar{x})$

coincides with the regular normal cone $\widehat{N}(X,\bar{x})$ to X at $\bar{x}$ [6], so that the GCQ at $\bar{x}$ may be rewritten as $\widehat{N}(X,\bar{x}) \subseteq L^\circ(g,h,\bar{x})$. The GCQ holds at $\bar{x}=0$ in Example 2.2.

Lemma 4.1. *A point $\bar{x}$ is a strong local minimizer of P if and only if $\bar{x} \in X$ and $-\nabla f(\bar{x}) \in \operatorname{int} \widehat{N}(X,\bar{x})$ hold.*

Proof. Since $T(X,\bar{x})$ is a closed cone, [6, Ex. 6.22] yields

$$\operatorname{int} \widehat{N}(X,\bar{x}) = \{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in T(X,\bar{x}) \setminus \{0\}\}. \tag{5}$$

Hence $-\nabla f(\bar{x}) \in \operatorname{int} \widehat{N}(X,\bar{x})$ is equivalent to $C_\leq(f,\bar{x}) \cap T(X,\bar{x}) = \{0\}$, and Lemma 2.1 yields the assertion. □

The GCQ at $\bar{x} \in X$ implies

$$\operatorname{int} \widehat{N}(X,\bar{x}) \subseteq \operatorname{int} L^\circ(g,h,\bar{x}). \tag{6}$$

Since also $L(X,\bar{x})$ is a closed cone, in analogy to (5) [6, Ex. 6.22] implies $\operatorname{int} L^\circ(g,h,\bar{x}) = \{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in L(g,h,\bar{x}) \setminus \{0\}\}$, and (6) is equivalent to

$$\{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in T(X,\bar{x}) \setminus \{0\}\} \subseteq \{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in L(g,h,\bar{x}) \setminus \{0\}\}. \tag{7}$$

Theorem 4.2. *Let the GCQ hold at $\bar{x} \in X$. Then $\bar{x}$ is a strong local minimizer of P if and only if it is a strong KKT point.*

Proof. The relation $T(X,\bar{x}) \subseteq L(g,h,\bar{x})$ and the combination of Lemma 2.1 with Lemma 3.4 yield that being a strong KKT point is sufficient for $\bar{x} \in X$ to be a strong local minimizer. For the proof of the reverse direction let $\bar{x}$ be a strong local minimizer. By Lemma 4.1 this is equivalent to $-\nabla f(\bar{x}) \in \{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in T(X,\bar{x}) \setminus \{0\}\}$. Since the GCQ at $\bar{x}$ implies (7) we conclude $-\nabla f(\bar{x}) \in \{v \in \mathbb{R}^n \mid v^\intercal d < 0 \, \forall d \in L(g,h,\bar{x}) \setminus \{0\}\}$. The latter is equivalent to $C_\leq(f,\bar{x}) \cap L(g,h,\bar{x}) = \{0\}$ and, by Lemma 3.4, to $\bar{x}$ being a strong KKT point. □

An analysis of the proof of Theorem 4.2 seems to indicate that in its assumption GCQ may be replaced by the weaker condition (6), that is, by an even weaker constraint qualification than the GCQ. However, (6) is not strictly weaker, but rather implies GCQ in the only relevant case for the proof of Theorem 4.2, so that both conditions are then equivalent. Indeed, in view of Lemma 4.1 we are only interested in the case $\operatorname{int} \widehat{N}(X,\bar{x}) \neq \emptyset$.

Proposition 4.3. *For $\operatorname{int} \widehat{N}(X,\bar{x}) \neq \emptyset$ the condition (6) implies the GCQ at $\bar{x}$.*

Proof. For $\operatorname{int} \widehat{N}(X,\bar{x}) \neq \emptyset$ the convex set $\widehat{N}(X,\bar{x})$ and, under (6), also $L^\circ(g,h,\bar{x})$ are full dimensional, so that their relative interiors coincide with their interiors. Therefore [5, Th. 6.3] implies $\operatorname{cl}\operatorname{int} \widehat{N}(X,\bar{x}) = \operatorname{cl} \widehat{N}(X,\bar{x}) = \widehat{N}(X,\bar{x})$ as well as $\operatorname{cl}\operatorname{int} L^\circ(g,h,\bar{x}) = \operatorname{cl} L^\circ(g,h,\bar{x}) = L^\circ(g,h,\bar{x})$, where the respective second identities follow from the closedness of polar cones. Since (6) yields $\operatorname{cl}\operatorname{int} \widehat{N}(X,\bar{x}) \subseteq \operatorname{cl}\operatorname{int} L^\circ(g,h,\bar{x})$, the GCQ follows. □

We remark that (6) is trivially fulfilled at any $\bar{x} \in X$ with $\operatorname{int} \widehat{N}(X, \bar{x}) = \emptyset$. In this case the GCQ may be violated at $\bar{x}$, as the example $X = \{x \in \mathbb{R}^2 \mid x_1^3 \leq 0\}$ with $\bar{x} = 0$ shows.

In view of Theorem 4.2 and Proposition 4.3 the GCQ at $\bar{x}$ may be the weakest condition under which a point $\bar{x} \in X$ is a strong KKT point for every at $\bar{x}$ continuously differentiable function f possessing $\bar{x}$ as a strong local minimizer on X. The following result verifies this.

Theorem 4.4. *The weakest condition under which a point $\bar{x} \in X$ is a strong KKT point for every at $\bar{x}$ continuously differentiable function f possessing $\bar{x}$ as a strong local minimizer on X, is the GCQ at $\bar{x}$.*

Proof. By Theorem 4.2 the GCQ at $\bar{x}$ is some condition under which, for every at $\bar{x}$ continuously differentiable function f possessing $\bar{x}$ as a strong local minimizer on X, $\bar{x}$ is also a strong KKT point. On the other hand, let $\bar{x}$ be a strong KKT point for every at $\bar{x}$ continuously differentiable function f possessing $\bar{x}$ as a strong local minimizer on X. We will show that then the GCQ $\widehat{N}(X, \bar{x}) \subseteq L^\circ(g, h, \bar{x})$ necessarily holds at $\bar{x}$.

In a first step we show that under the current assumption the GCQ trivially holds at $\bar{x}$ in the case $\operatorname{int} \widehat{N}(X, \bar{x}) = \emptyset$, since then no continuously differentiable function f can possess $\bar{x}$ as a strong local minimizer. Indeed, Lemma 4.1 would then imply $-\nabla f(\bar{x}) \in \operatorname{int} \widehat{N}(X, \bar{x})$, in contradiction to $\operatorname{int} \widehat{N}(X, \bar{x}) = \emptyset$.

In the remainder of the proof let $\operatorname{int} \widehat{N}(X, \bar{x}) \neq \emptyset$. Then by Proposition 4.3 it is sufficient to show (6) for the proof of GCQ at $\bar{x}$. Indeed, choose $v \in \operatorname{int} \widehat{N}(X, \bar{x})$. We need to show that $v^\intercal d < 0$ holds for all $d \in L(g, h, \bar{x}) \setminus \{0\}$. By Lemma 4.1 the linear function $f(x) = -v^\intercal x$ possesses $\bar{x}$ as a strong local minimizer on X. By assumption $\bar{x}$ is also a strong KKT point of f on X, that is, $\operatorname{rank}(\nabla f(\bar{x}), \nabla g_A(\bar{x}), \nabla h(\bar{x})) = n$ holds and there exist $\lambda > 0$, $\mu > 0$ and ν with

$$v = \nabla g_A(\bar{x})\lambda + \nabla h(\bar{x})\nu.$$

For all $d \in L(g, h, \bar{x}) \setminus \{0\}$ this implies

$$v^\intercal d = \lambda^\intercal \nabla g_A(\bar{x})^\intercal d + \nu^\intercal \nabla h(\bar{x})^\intercal d \leq 0.$$

Moreover, in the case $v^\intercal d = 0$ we would obtain

$$0 = v^\intercal d = \lambda^\intercal \nabla g_A(\bar{x})^\intercal d + \nu^\intercal \nabla h(\bar{x})^\intercal d$$

which, in view of $\lambda > 0$, is only possible for $\nabla g_A(\bar{x})^\intercal d = 0$. Therefore we arrive at $0 = v^\intercal d = -\nabla f(\bar{x})^\intercal d$, $\nabla g_A(\bar{x})^\intercal d = 0$ and $\nabla h(\bar{x})^\intercal d = 0$, that is, $d^\intercal(\nabla f(\bar{x}), \nabla g_A(\bar{x}), \nabla h(\bar{x})) = 0$. The rank condition (3) implies $d = 0$, which contradicts the choice $d \neq 0$. We have thus shown $v^\intercal d < 0$ for all $d \in L(g, h, \bar{x}) \setminus \{0\}$. □

References

[1] G. Giorgi, B. Jiménez, and V. Novo. “A Note on First-Order Sufficient Optimality Conditions for Pareto Problems”. In: *Numerical Functional Analysis and Optimization* 29.9-10 (2008), pp. 1108–1113.

[2] F.J. Gould and J.W. Tolle. "A necessary and sufficient qualification for constrained optimization". In: *SIAM Journal on Applied Mathematics* 20.2 (1971), pp. 164–172.

[3] B. Jiménez and V. Novo. "First and second order sufficient conditions for strict minimality in nonsmooth vector optimization". In: *Journal of Mathematical Analysis and Applications* 284.2 (2003), pp. 496–510.

[4] O.L. Mangasarian. *Nonlinear Programming*. SIAM, 1994.

[5] R.T. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[6] R.T. Rockafellar and R.J. Wets. *Variational Analysis*. Springer Science & Business Media, 2009.

[7] G. Still and M. Streng. "Optimality conditions in smooth nonlinear programming". In: *Journal of Optimization Theory and Applications* 90.3 (1996), pp. 483–515.